

Surface Wind Regionalization in Complex Terrain

P. A. JIMÉNEZ,* J. F. GONZÁLEZ-ROUCO,⁺ J. P. MONTÁVEZ,[#] J. NAVARRO,[@]
E. GARCÍA-BUSTAMANTE,* AND F. VALERO⁺

**División de Energías Renovables, CIEMAT, and Departamento de Astrofísica y Ciencias de la Atmósfera, Universidad Complutense de Madrid, Madrid, Spain*

⁺ Departamento de Astrofísica y Ciencias de la Atmósfera, Universidad Complutense de Madrid, Madrid, Spain

[#] Departamento de Física, Universidad de Murcia, Murcia, Spain

[@] División de Energías Renovables, CIEMAT, Madrid, Spain

(Manuscript received 25 April 2006, in final form 9 May 2007)

ABSTRACT

Daily wind variability in the Comunidad Foral de Navarra in northern Spain was studied using wind observations at 35 locations to derive subregions with homogeneous temporal variability. Two different methodologies based on principal component analysis were used to regionalize: 1) cluster analysis and 2) the rotation of the selected principal components. Both methodologies produce similar results and lead to regions that are in general agreement with the topographic features of the terrain. The meridional wind variability is similar in all subregions, whereas zonal wind variability is responsible for differences between them. The spectral analysis of wind variability within each subregion reveals a dominant annual cycle and the varying presence of higher-frequency contributions in the subregions. The valley subregions tend to present more variability at high frequencies than do higher-altitude sites. Last, the influence of large-scale dynamics on regional wind variability is explored by studying connections between wind in each subregion and sea level pressure fields. The results of this work contribute to the characterization of wind variability in a complex terrain region and constitute a framework for the validation of mesoscale model wind simulations over the region.

1. Introduction

Understanding and forecasting regional wind variability is relevant for a wide variety of phenomena, for example, the transport and dispersion of pollutants along an area, planning and decision making within situations of risk assessment related to the occurrence of extreme events such as forest fires or structural damage, and power forecasting in wind farms. The latter is becoming an issue with increasing demand for establishing national policies in view of the recent developments of renewable energy technologies.

Regional variability is controlled by the interaction of large-scale dynamics and orography. Circulation in the free atmosphere is governed by gradients between the large pressure systems. In the lower troposphere, the topography gains importance, generating a dynamical

forcing that modifies the direction and intensity of winds resulting from channeling, forced ascents, and barrier effects (Whiteman 2000). These dynamically driven circulations show variations within the frequencies of a few days. In addition, differential heating and cooling of the soil drive local thermal circulations (Blumen 1990), which depend on the temperature differences along the valley axis or the mountain plains systems, leading to diurnal frequency variations in response to the solar heating (Whiteman 2000).

Wind variability, therefore, is more complicated in complex terrain regions where dynamically and thermally driven wind systems and their interactions generate a wide variety of flow patterns (McGowan and Sturman 1996). Furthermore, thermally driven winds at the large scale can overwhelm those at the local scale as shown by Stewart et al. (2002) in their study of four regions in western United States. The relationship between the synoptic scale and the flow within a valley was studied by Whiteman and Doran (1993), who suggested that the thermal forcing occurs when the large-scale dynamical forcing is weak. Because the dynamical

Corresponding author address: J. Fidel González-Rouco, Facultad de Ciencias Físicas, Universidad Complutense de Madrid, Avd. Complutense s/n, Madrid 28040, Spain.
E-mail: fidelgr@fis.ucm.es

cally driven circulations are controlled by the synoptic-scale motions and the thermally driven circulations become more relevant when these motions are weak, synoptic-scale motions either directly or indirectly control the circulations over the surface. Their typical synoptic variability of a few days makes it suitable to undertake a daily wind study.

The complicated circulations over complex terrain regions can be better understood by dividing the area into a small number of internally homogeneous subregions, by means of a wind regionalization. One strategy to obtain areas of distinct regional behavior is to use eigenvector techniques. These techniques have been successfully applied in a variety of case studies and for several variables (Dyer 1975; Barring 1988; Stooksbury and Michaels 1991; Bonell and Sumner 1992; Fovell and Fovell 1993; Comrie and Glenn 1998; Romero et al. 1999a). However, the potential of the regionalization approach seldom has been explored with wind-related variables (Cheng 1998). Regionalization of climate parameters over a specific area contributes to the improvement of the understanding of the spatial and temporal variability of the climate variable, and also offers a framework for the validation of mesoscale model simulations over the region (Romero et al. 1999a; Sotillo et al. 2003). Mesoscale models have become a standard tool for providing simulations and forecasts of airflows in complex terrain regions, favored by recent increases in computational power and accessibility of analysis and forecast grids (Mass and Kuo 1998). An increased understanding of wind variability provides a better validation of mesoscale simulations (Rife et al. 2004), and consequently the improvement of their accuracy. Regionalization allows for the establishment of subregions of distinct behavior in observations that should be reproduced by mesoscale models, thus offering the possibility for designing model configurations and test sensitivities either to initial and boundary conditions or to different physical parameterizations. The model evaluation at various identified subregions, instead of a point-to-point approach, that is, at specific grid points or sites, allows the influence of local effects that are not usually modeled to be mitigated (von Storch 1995).

This paper addresses the problem of wind regionalization in a complex terrain region at daily time scales. Two different methods based on principal component analysis (PCA) are used. The first method carries out cluster analysis (CA) of the most important PCA modes (Romero et al. 1999b), which allows for the identification of homogeneous wind climate variability groups, while the second method makes use of the rotation of selected principal components (White et al.

1991). The temporal variability of wind in each subregion is investigated by analyzing the spectra of the selected principal components. The Comunidad Foral de Navarra (CFN) region in northern Spain was selected as the case study (Fig. 1) for this purpose. Its complex topography and strong wind conditions, which have resulted in increases of wind farm facilities in recent times, make it an interesting place to study wind variability. The CFN region presents a complicated, highly variable topography with numerous valleys and mountain ridges that produce an interesting degree of interaction with large-scale dynamics (Fig. 1), thus constituting a useful case both for improving our understanding of wind variability at regional scales and for future application in validation assessments of model performance.

This paper is organized as follows: Sections 2 and 3 briefly describe the datasets and the multivariate methods used in the regionalization process. Section 4 describes results obtained from the CA and the PCA rotation-based approaches and further illustrates aspects of the temporal wind variability in each region. Conclusions and a discussion are presented in section 5.

2. Data

The location of the CFN in the north of the Iberian Peninsula is highlighted in Fig. 1. The orography of the region shows a variety of rich features broadly limited by two large mountain systems: the Iberic System in the south of the CFN and the Pyrenees in the north, which merge westward with the last foothills of the Cantabrian Mountains. Between them, the Ebro Valley crosses the region from northwest to southeast toward the Mediterranean. A closer look at the CFN (see enlarged area in Fig. 1) reveals a complex array of smaller mountain systems and valleys. The north is dominated by the Pyrenees, with the large Bidasoa mountain lines or the smaller sierras of Abodi, Uztarroiz, San Miguel, Zariquieta, and Leyre. To simplify, the latter will be referred to as the *northern mountains*. The western and northwestern boundaries are outflanked by the Aralar mountain lines as well as Urbasa, Santiago, and Andía, which will be herein labeled as *western mountains*, unless treated specifically. The center and eastern side of the CFN is punctuated by the mountains systems of Izco, San Pedro, and the Ujué peak, which will be labeled as the *eastern mountains*. Last, the south of the CFN is dominated by the lower lands of the Ebro Valley. Virtually parallel to the Ebro Valley, smaller valleys line up in the northwest–southeast direction south of the northern mountains. Similarly, west and east of the eastern mountain group, several mountain systems

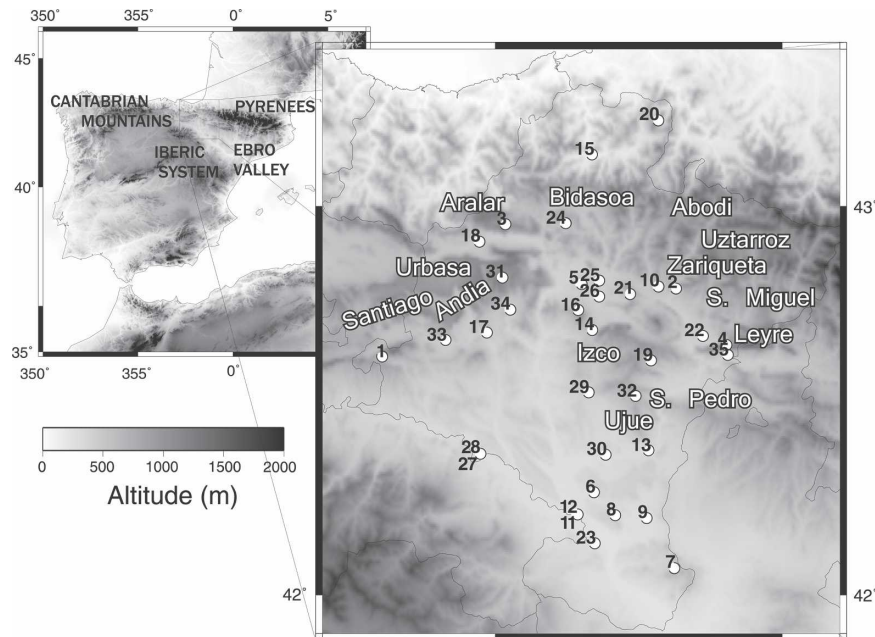


FIG. 1. The location of wind stations within the CFN. Shading represents altitude, circles are the measurement sites, and the thin lines highlight political boundaries. See Table 1 for specific station descriptions. (left) The most important geographical features around the CFN are highlighted, and (right) some regional details of the CFN are listed.

and valleys seem to favor wind channeling in the north–west–southeast and north–south directions. It will be illustrated that this is a major feature of the variability of the wind field within the region.

The dataset spans the period from 1 January 1992 to 30 September 2002. The best 35 stations with the best-quality wind measurements from the meteorological network of the CFN were selected for this study (Fig. 1, Table 1). Observations of wind speed and direction were recorded at a height of 10 m above ground level, with the exception of seven stations in which the measurements were taken at 2 m, and with a time resolution of 10 min (Table 1).

These initial data were quality controlled by applying various tests that are similar to those employed in other quality control analysis (Meek and Hatfield 1994; De-Gaetano 1997; Graybeal 2006). In particular, several tests were applied to identify and eliminate errors of sensor or data manipulation (repetition, unrealistically high or low values, etc.), to invalidate abnormally high- or low-variability periods in order to ensure temporal consistency, as well as to assess the long-term variability of the time series (Jiménez et al. 2007, unpublished manuscript). The resulting data were subsequently transformed to zonal and meridional wind components and averaged daily to perform the wind regionalization. Some of the stations were installed after 1992 and have either less or no observations during the first years of

the record. The existence of missing data, and thus of uneven time intervals, can potentially produce adverse effects in the calculation of principal components. This can be partially mitigated by interpolating the data into a grid (Wilks 1995). However, grid interpolation of wind data can be particularly problematic in complex terrain regions where exposure, orographic features, and altitude are usually more important than distance (Kaufmann and Weber 1998; Steinacker et al. 2006). In addition, interpolation does not incorporate new information to the dataset unless more variables or sites are considered. An alternative possibility is to undertake a careful *pairwise* treatment of missing values, as suggested by Barring (1988). Ludwig et al. (2004) applied PCA both to grid-interpolated and unevenly spaced station data for the same region and obtained similar results; other examples of PCA with unevenly spaced data are those of Bonell and Sumner (1992) and Comrie and Glenn (1998).

This work makes use of this last approach by applying principal component analysis to a subset of data with fewer missing data in its daily fields in order to secure a more robust estimation of eigenvectors. The complete dataset is posteriorly used to calculate an extended version of the principal components, which allows for study of the variability in each region at longer time scales (see section 4c). The subset of data is selected, including only the daily fields with more than

TABLE 1. The codes of the stations as in Fig. 1, name, longitude, latitude, altitude, and height of the sensor.

No.	Name	Lon (°)	Lat (°)	Alt (m)	Sensor height (m)
1	Aguilar de Codés	−2.394	42.614	731	10
2	Aoiz-Agoitz	−1.369	42.792	530	10
3	Aralar	−1.963	42.954	1393	10
4	Arangoiti	−1.194	42.646	1353	10
5	Arazuri	−1.702	42.801	396	2
6	Bardenas-barranco salado	−1.654	42.265	300	2
7	Bardenas-loma negra	−1.375	42.071	646	10
8	Bardenas-Nstra Sra. del Yugo	−1.582	42.206	472	10
9	Bardenas-polígono de tiro	−1.473	42.200	295	10
10	Beortegi	−1.434	42.796	580	10
11	Cadreira-Riegos	−1.717	42.209	268	2
12	Cadreira-INM	−1.710	42.208	268	10
13	Carcastillo	−1.463	42.372	340	10
14	Carrascal	−1.660	42.683	560	10
15	Doneztebe	−1.660	43.132	138	10
16	Perdón	−1.709	42.733	1024	10
17	Estella-Lizarra	−2.028	42.676	480	10
18	Etxarri-Aranatz	−2.057	42.910	507	10
19	Getadar	−1.457	42.605	710	10
20	Gorramendi	−1.432	43.220	1071	10
21	Ilundain	−1.529	42.777	542	10
22	Lumbier-Ilumberri	−1.275	42.668	484	2
23	Montes del Cierzo	−1.652	42.133	310	10
24	Oskotz	−1.756	42.956	562	10
25	Pamplona-Larrabide	−1.638	42.810	450	10
26	Pamplona-Noain	−1.639	42.769	461	10
27	Sartaguda-Riegos	−2.050	42.363	310	2
28	Sartaguda-INM	−2.051	42.366	310	10
29	Tafalla	−1.676	42.522	415	10
30	Traibuenas	−1.614	42.363	312	2
31	Trinidad de Iturgoien	−1.975	42.819	1222	10
32	Ujué	−1.510	42.513	826	10
33	Valdega	−2.172	42.657	469	2
34	Villanueva del Yerri	−1.949	42.736	498	10
35	Yesa	−1.190	42.618	489	10

80% of the site observations available (i.e., more than 28 stations), which reduces the number of available daily fields to 947 but ensures a homogeneous representation of all stations during the time steps for the calculation of eigenvectors. A further selection of fields was additionally made considering the resulting monthly distribution of available daily fields (Fig. 2). The irregular spread of data along the year could potentially stress the prevailing circulations of the months with more available daily fields. This potential undesirable effect was weakened by imposing an upper limit on the retained number of daily fields for each month, thus

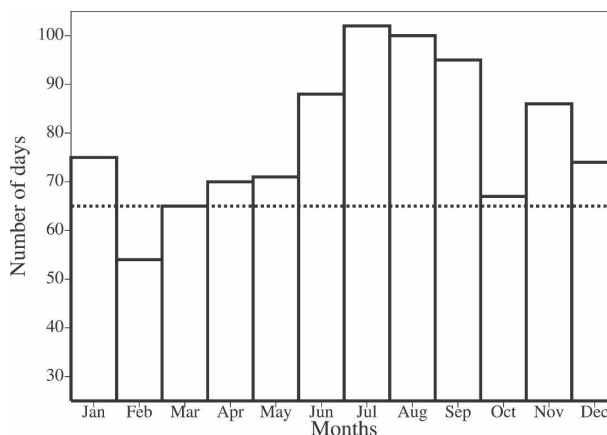


FIG. 2. The monthly distribution of the number of days with more than 80% of the wind observations available (boxes). The dashed line represents the final homogeneous distribution after discarding the days with larger amounts of missing values from each month (except for February, which has only 54 available days).

achieving a more homogeneous distribution. This threshold was established, in this case selecting the best-quality 65 daily cases for each month (dashed line in Fig. 2), with the only exception being February, which could only accumulate 54 daily fields. The final subset containing a total of 769 daily wind measurement fields was employed for the wind regionalization. Because the quality of the original dataset was progressively improved with time, the latter subset was mostly concentrated in the recent years, spanning the period from October 1999 to September 2002.

3. Methodologies

A vectorial PCA (Kaihatu et al. 1998; Ludwig et al. 2004) was applied to the correlation matrix of the available zonal and meridional time series (S mode; see Richman 1986) as a first step in both regionalization methods. The use of the correlation matrix, instead of the covariance matrix, allows for the comparison of sites with different ranges of wind variability. PCA can be applied in a scalar approach to each wind component or in a vectorial one, maximizing the joint variance of both components. Klink and Willmott (1989) favor the use of the vectorial approach performed herein as a more general and useful way to explain wind variability by exploiting the shared variance between the zonal and meridional components.

PCA is a standard methodology, and the reader is referred to textbooks for more information (Preisendorfer 1988; von Storch and Zwiers 1999). A brief definition is presented in Eq. (1),

$$[u(x, t), v(x, t)] = [\bar{u}(x), \bar{v}(x)] + \sum_{i=1}^{2N} a_i(t)[u_i(x), v_i(x)], \quad (1)$$

where the elements on the left-hand side are the zonal $u(x, t)$ and meridional $v(x, t)$ wind data at time t and at each of the N sites ($x = 1, 2, \dots, N$). Here, \bar{u} and \bar{v} represent the time average of the zonal and meridional wind components. The second term on the right-hand side involves the $i = 1, \dots, 2N$ principal modes through the product of the time series of scores (principal components) $a_i(t)$, and the eigenvector $[u_i(x), v_i(x)]$. The selection of the number of principal modes to retain M was made with the classical scree test of Cattell (1966).

a. Cluster analysis methodology

The CA of the retained PCA modes allows for obtaining groups of stations with similar loads. The target of any CA is to form groups with large similarities between objects inside a cluster and small similarities between objects of different clusters. Hence, the first step to perform a CA is to define a similarity measure between the objects to be grouped. In this analysis the objects are the stations, and the measure of similarity between two given stations A and B is defined by

$$d_{AB} = \frac{1}{M} \sum_{i=1}^M \{[u_i(x_A) - u_i(x_B)]^2 + [v_i(x_A) - v_i(x_B)]^2\}^{1/2}, \quad (2)$$

where i denotes the principal mode and u and v are the components of the vectorial loads as in Eq. (1). Once the similarity measure has been defined, a clustering procedure must be chosen. Hierarchical and nonhierarchical algorithms are standard tools for this purpose (Anderberg 1973). A two-step cluster procedure is frequently used (Stooksbury and Michaels 1991; Kaufmann and Weber 1996; Kaufmann and Whiteman 1999) because of its advantages against the single-step cluster algorithms (Milligan 1980), and thus it was the method chosen for this study. In the first step a hierarchical algorithm finds the groups that serve as an initial state for a second phase in which a nonhierarchical scheme reorders the objects. Weber and Kaufmann (1995) compared several hierarchical algorithms for a wind field classification finding advantages of the complete linkage algorithm (CLA; Johnson 1967), and so it was the algorithm selected for the first step. This algorithm defines the similarity between two clusters as the largest value of the distance [Eq. (2)] between any possible pair of objects, with one from each cluster. With this definition, CLA merges two clusters only if the distance

between the most dissimilar objects does not exceed a certain threshold. In this way, the method ensures broad similarity of the elements within a cluster and dissimilarity of these against those of other clusters. To decide on the appropriate number of clusters for formation, the value of the similarity measure at which the two most similar clusters are merged at each step is screened; a large change in the value of the distance before and after merging two potentially similar groups means that two very different clusters have been merged and can be used as an indication to stop the algorithm in the previous step.

On the second CA step, a method similar to the nonhierarchical k means procedure (Kaufmann and Weber 1996) is employed. This algorithm calculates the similarity, based in Eq. (2), between each station and a reference centroid representative of the cluster to which the station has been previously assigned. Initial cluster assignment is made in the previous step with the CLA, and centroids are calculated as the average of all of the individuals within each initial group. After this, distances according to Eq. (2) are calculated on the basis of the loadings of each target station and those of the centroid. This allows for a new reassignment through which stations can be moved to a different group, which presents minimum distance between its centroid and the target station. Once the procedure has been applied to each site, new reassignment steps can be undertaken in an iterative manner until stability is attained and no station is virtually relocated in a different group.

b. Rotation of principal components methodology

An alternative regionalization approach has been explored and is based on the rotation of the selected principal modes to obtain the wind regions. The aim of the rotation is to produce "simple structure" in which the variables are as close as possible to a hyperplane of at least one principal mode (Richman 1986). With the simple structure, the loading map of each principal mode weighs a different subregion.

Theoretically, each observational site should present a high load in just one rotated loading map and null loads in the rest (perfect simple structure). This would imply that each map defines a completely different subregion. Actually, the loads are not null and it is necessary to define a *critical* threshold value to define the subregions: only those sites with loads higher than the critical value will belong to a specific subregion. Because a vectorial PCA is adopted the loads are vectors and the critical value is defined based on the value of its module.

Rotations can be either orthogonal or oblique. There is considerable discussion on the benefits and disadvantages of each type of rotation. Some authors find very similar groups with both methods (Gregory 1975), while others conclude that the oblique rotations produce more stable results and are superior to the orthogonal rotations (White et al. 1991). Varimax (orthogonal) and oblimin (oblique) rotation techniques were tested in this work and little difference was found for the case of this study. The first technique was finally selected because of its simplicity and property of preserving the orthogonality of the eigenvectors after the rotation.

This regionalization method allows for one station to belong to more than one subregion because the regions derived from each PCA mode can overlap, contrary to the CA method, which generates a hard regionalization in which each station belongs to only one subregion. Furthermore, because this second regionalization method assigns one principal mode to each subregion, it makes it possible to analyze the wind variability of each subregion by calculating the spectra of its corresponding time series of scores. The series of scores are not continuous in time because the input data presented missing values, and only daily fields with a high percentage of available measurements were selected (see section 2). Standard autocovariance Fourier transform spectrum analysis (Bloomfield 1976) can find difficulties in its application in cases of large amounts of missing values. This case can be treated as one of the irregular samplings of data (Belserene 1988), and therefore spectra are calculated herein with an alternative approach that does not require equidistant sampling (Deeming 1975). The spectral estimate is comparable to a normalized periodogram and can be interpreted as such, but no limitation is imposed on the regular or irregular character of the sampling when calculating the discrete Fourier transform. A spectral window that contains the time scales of interest is selected and the spectral estimate for a set of trial frequencies is obtained, which in this approach will be regularly distributed over the spectral window. For details and discussion on this approach, the reader is referred to Deeming (1975) and Belserene (1988).

4. Results

As a preliminary inspection of the wind variability over the region, the average and standard deviation fields of the wind speed module are displayed in Fig. 3. The spatial patterns of both variables suggest a linear relation between them, with sites showing higher wind averages also presenting more variability. This is typical

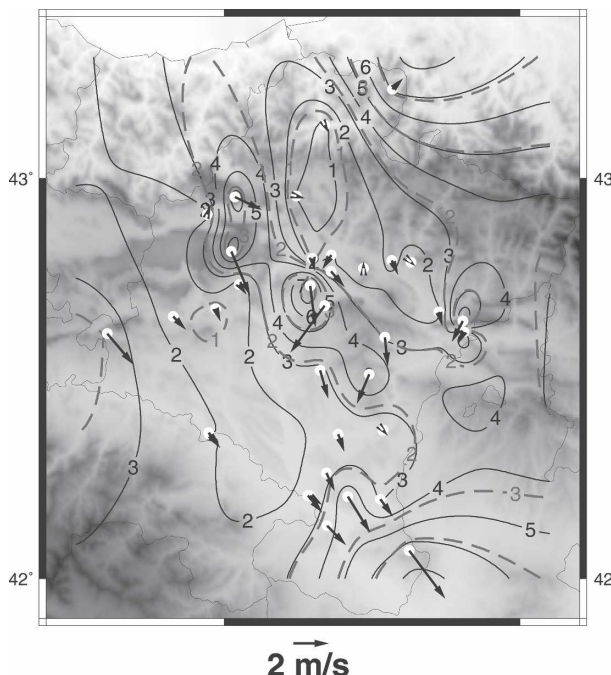


FIG. 3. Mean wind speed field (solid lines), its standard deviation (dashed lines), and mean zonal and meridional wind components (vectors) of the selected daily wind fields. The tail of vectors is placed at the observational sites.

of positively defined variables, like precipitation (Xoplaki et al. 2004), in which increases in mean values lead to wider probability distributions (higher variability). For the case under consideration, Fig. 4a illustrates this relationship (correlation $r = 0.98$). The sites with the strongest wind are the mountain stations and some of the stations in the Ebro Valley (Fig. 3). Many of the windy sites are also the highest in altitude (Table 1), suggesting a relation between altitude and wind speed. This can be better observed by displaying the dispersion diagram of both variables (Fig. 4b), which shows a correlation value of 0.78.

The mean wind vectors displayed in Fig. 3 are calculated from averages of the zonal and meridional wind components. The mean flow is southeastward and is channeled from the northern valleys to the Ebro Valley through the north–south passages around the eastern mountains. The meridional component presents, in general, higher variability than the zonal component (Fig. 4c). The correlation between both components ($r = 0.76$) indicates some linear relationship between them. This can be argued to be a reasonable feature from the point of view of the conservation of horizontal momentum. Assuming that the loss of energy from surface friction is small and the loss of horizontal momentum resulting from vertical ascents or descents represent a small fraction of horizontal momentum on daily

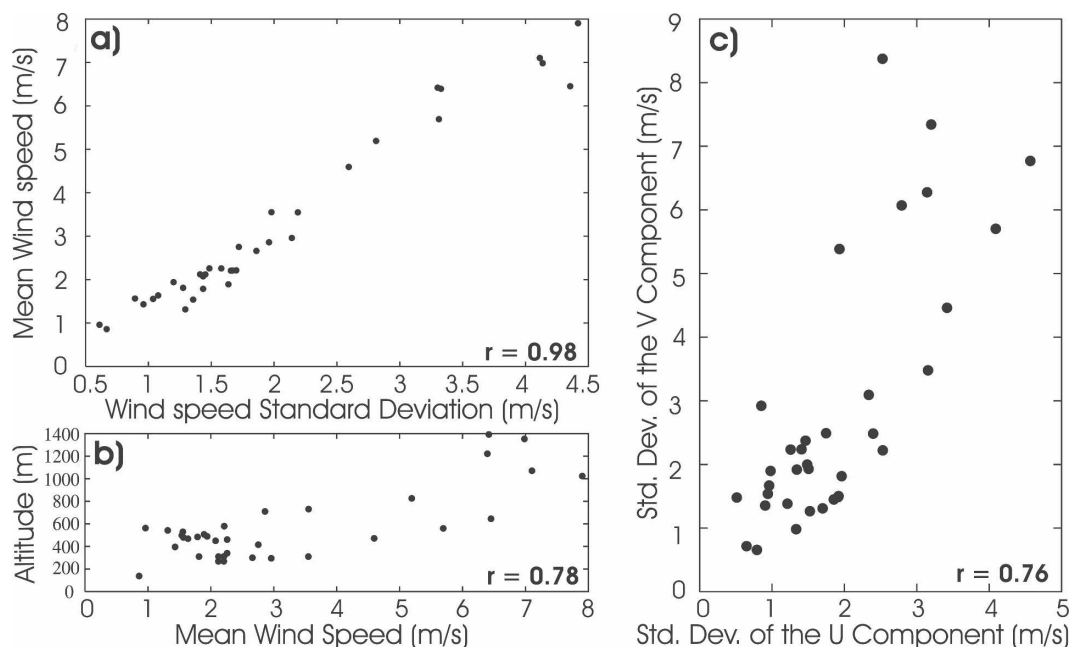


FIG. 4. (a) Mean daily wind speed vs its standard deviation for the selected days in the 35 sites, (b) altitude vs mean wind speed, and (c) standard deviation of the v wind component vs standard deviation of the u wind component.

time scales, changes in the zonal (meridional) component resulting from the interaction of the wind with orographic obstacles will translate into the transfer of momentum to the meridional (zonal) component. Thus, changes in the zonal (meridional) wind component will often be related to changes in the meridional (zonal) component, and ultimately sites showing higher variability in one component should be expected to have a higher variability in the other component also. This relation is an indication of common variability and supports the joint treatment of both components in the application of the vectorial PCA in the analysis of the wind variability instead of performing a separate analysis on each variable component. As for the higher meridional than zonal variability, this is related (not shown) to the channeling of the flow between the large mountain systems in northern Spain (the Cantabrian Mountains and the Pyrenees; see Fig. 1), and at a more regional scale within the CFN. Here, channeling is favored along the northern valleys and the Ebro Valley, with a northwest–southeast orientation, and particularly around the eastern mountains systems, with a north–south orientation. This behavior will be further illustrated with the results of the PCA analysis in the following sections.

a. Regionalization using cluster analysis

The explained variance of the leading PCA modes of the wind field is shown in Fig. 5. There are breaks of the

slope at modes three and five, as well as a less clear one at mode nine, which according to the scree test of Cattell (1966), are reasonable numbers of modes to retain. The retention of three of them might not be enough to

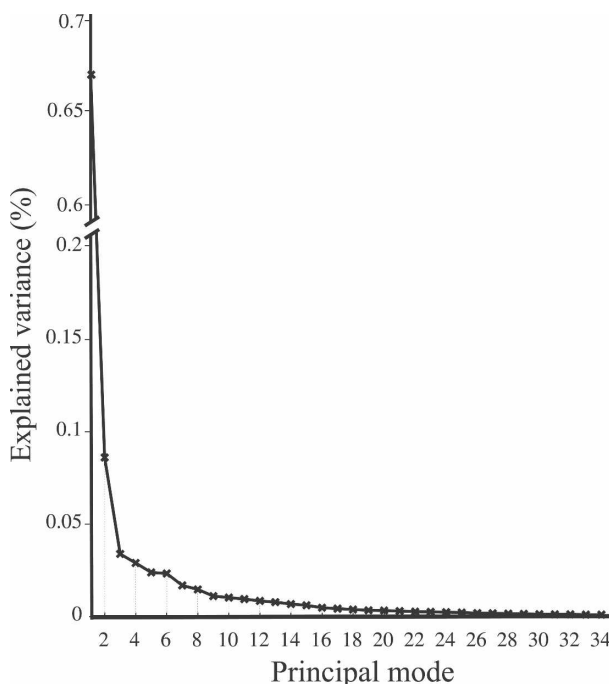


FIG. 5. The explained variance of the PCA modes. Notice the break in the vertical scale.

adequately group the stations, because in such cases the similarity measure (2) would be calculated with the contributions of only three terms, whereas nine modes could introduce noise into the classification because of the little variance explained by the higher-order modes. Therefore, five principal modes, which accumulate 84.1% of the variance in the data, were retained. The loading maps (eigenvectors) of these principal modes are displayed in Fig. 6. These maps represent flow directions and can be interpreted in their positive phase as displayed either in Fig. 6 or in the opposite sign of the mode (corresponding to a negative sign in their principal component). The first mode explains two-thirds of the variance (66.8%) and is well organized (Fig. 6a): the vectors are aligned along the valley axis, indicating that the dominant physical process is the topographical channeling. In the positive (negative) phase the main flow has a southeast–northwest (northwest–southeast) direction, while some mountain sites seem to show a certain decoupling of the flow with respect to the valley circulations, presenting a meridional direction with a northward (southward) orientation. The second principal mode explains 8.6% of the variance and presents strong eastward (westward) flows at the highest locations mainly, and weaker southwest–northeast (northeast–southwest) flows at the rest of the sites (Fig. 6b). Physically, this can be interpreted as the influence of the synoptic-scale flows, which is more intense at the higher sites. Therefore, this mode also reveals a decoupling of the flow between mountain and valley circulations, noted before. The third principal mode explains 3.4% of the variance. It shows activity in the center of the region, with zonal directions in the flows but with opposite orientations at sites separated by relatively small distances that could be associated with the recirculation or local behavior of the flow (Fig. 6c). The fourth and fifth modes explain 2.9% and 2.4% of the variance, respectively, and show relations between a few stations of each mode (Figs. 6d,e). The low percentage of variance explained by these last modes can be related to the limited size of their area of influence relative to that of the three first modes. The improved knowledge of variability within this confined area could stem, in the future, from higher spatial sampling. This arrangement of explained variance is similar to that found by other authors in regions of comparable size (Hardy and Walton 1978; Green et al. 1992; Ludwig et al. 2004).

The first step of this CA regionalization is to group together the stations with similar loads (Fig. 6) using the hierarchical CLA in order to decide the number of subregions to be formed. This was done by displaying the sequence of distance measures at which the clusters

were merged in each step (Fig. 7). Because the CLA merges the two most similar clusters at each step, a large jump in the sequence means that two very different clusters have been merged, and indicates the convenience of stopping the algorithm just before this happens. Steps appear at six and nine clusters (Fig. 7), and therefore they are a reasonable number of subregions to form. The regionalization of nine subregions forms many small groups with only one or two stations (Fig. 8a); hence, a six-cluster regionalization was selected (Fig. 8b). After the CLA, the reordering of the stations in the six selected subregions is performed by the nonhierarchical algorithm and provides the final wind regionalization (Fig. 8c). This reordering only changes the location of station number 32 (Fig. 1), which was assigned to the first subregion by the CLA (Fig. 8b) and now belongs to the third wind region (Fig. 8c). The similar clustering of the stations obtained in the two steps of the CA methodology grants robustness to the proposed regionalization. The Ebro Valley stations form the first subregion (label 1 in Fig. 8c), the narrow northern valleys are the second subregion (label 2 in Fig. 8c), the high mountain stations are the third subregion (label 3 in Fig. 8c), and the rest of the subregions are small groups with a north-to-south orientation in the center of the CFN (labels 4, 5, and 6 in Fig. 8c). Conceptually, this seems to suggest that there are three well-defined regions, and a fourth one, the central north-to-south area that groups the clusters with the lowest numbers of stations.

b. Regionalization using rotated principal modes

This method rotates the retained PCA modes to perform the regionalization. Subregions are then formed retaining just the highest loads of each loading map. This is done by defining a critical threshold value, and only those sites with loads exceeding it will belong to the subregion. Hence, one subregion is created for each rotated mode; the method allows for the formation of overlapping subregions in opposition to the CA method, which generates a hard regionalization without possible overlap. The retention of five modes, as was done with the regionalization using CA, results in a strong overlapping between the formed subregions and suggests that five wind regions appear to be too many. Hence, only four principal modes were employed for this regionalization method (81.7% of the variance). As was mentioned above, rotation tends to form *simple structures* (Richman 1986) in which the variables are as close as possible to the hyperplane of at least one principal mode. The degree of the simple structure can be visualized displaying the loads of one principal mode

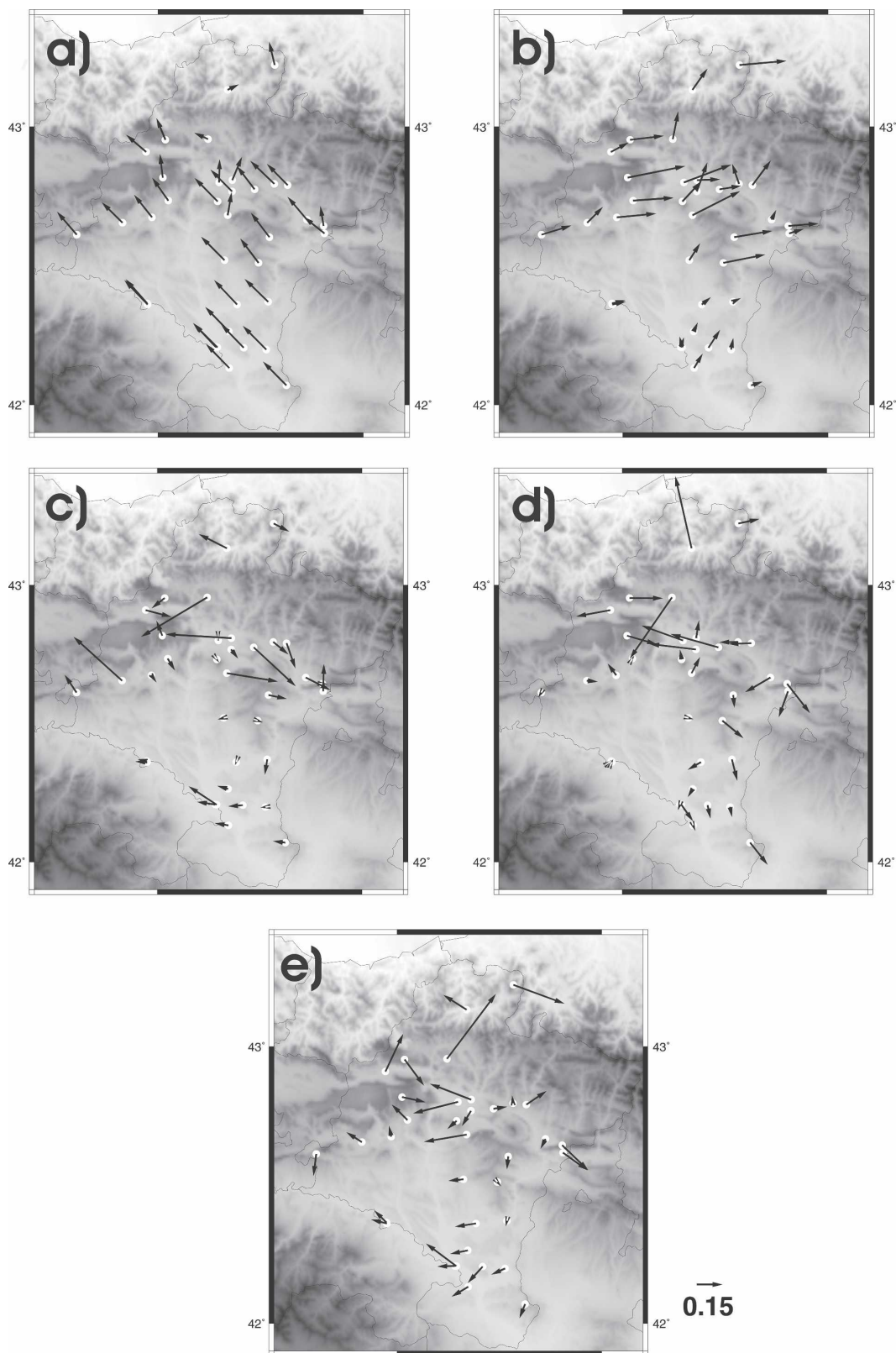


FIG. 6. Loading maps of the first five principal modes are shown. The tails of the vectors are placed at the observational sites. Explained variances for each mode are (a) 66.8%, (b) 8.6%, (c) 3.4%, (d) 2.9%, and (e) 2.4%.

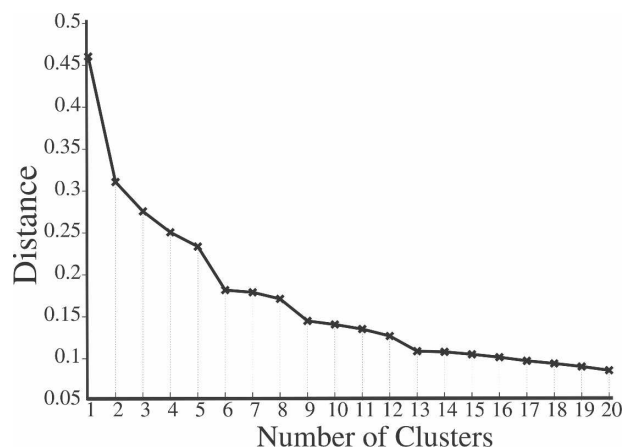


FIG. 7. The distance at which the last two clusters are merged against the number of clusters formed is shown. The larger jump at 6 and 9 suggests that these numbers of clusters should be retained.

against those of another mode. Figure 9 shows an example of this for mode 1 versus 2 and 4, illustrating that values overall tend to be closer to the axis in the rotated case. The effect of the varimax rotation is apparent in the turning of the dispersion diagram of mode 1 versus that of mode 4.

The loading maps of the four rotated principal modes in their positive phase are displayed in Fig. 10. When compared with the unrotated principal modes (Fig. 6), they show a clearer physical interpretation in some aspects. The first and second rotated principal modes (Figs. 10a,b) present similar patterns to those of the first two modes (Figs. 6a,b). The first pattern presents in its positive (negative) phase the southeast–northwest (northwest–southeast) channeling along the Ebro Valley, as in the unrotated case, and some minor differences in the representation of its influence in the area

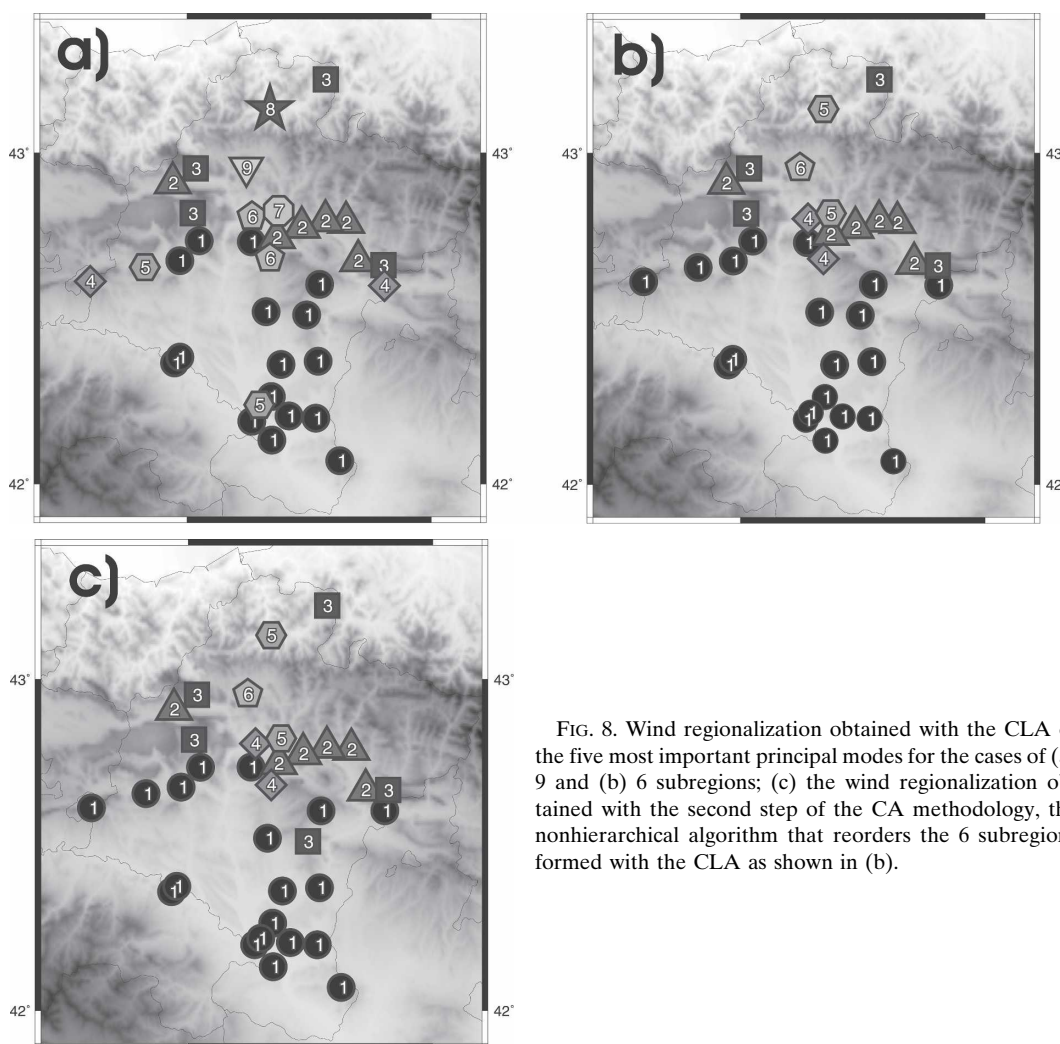


FIG. 8. Wind regionalization obtained with the CLA of the five most important principal modes for the cases of (a) 9 and (b) 6 subregions; (c) the wind regionalization obtained with the second step of the CA methodology, the nonhierarchical algorithm that reorders the 6 subregions formed with the CLA as shown in (b).

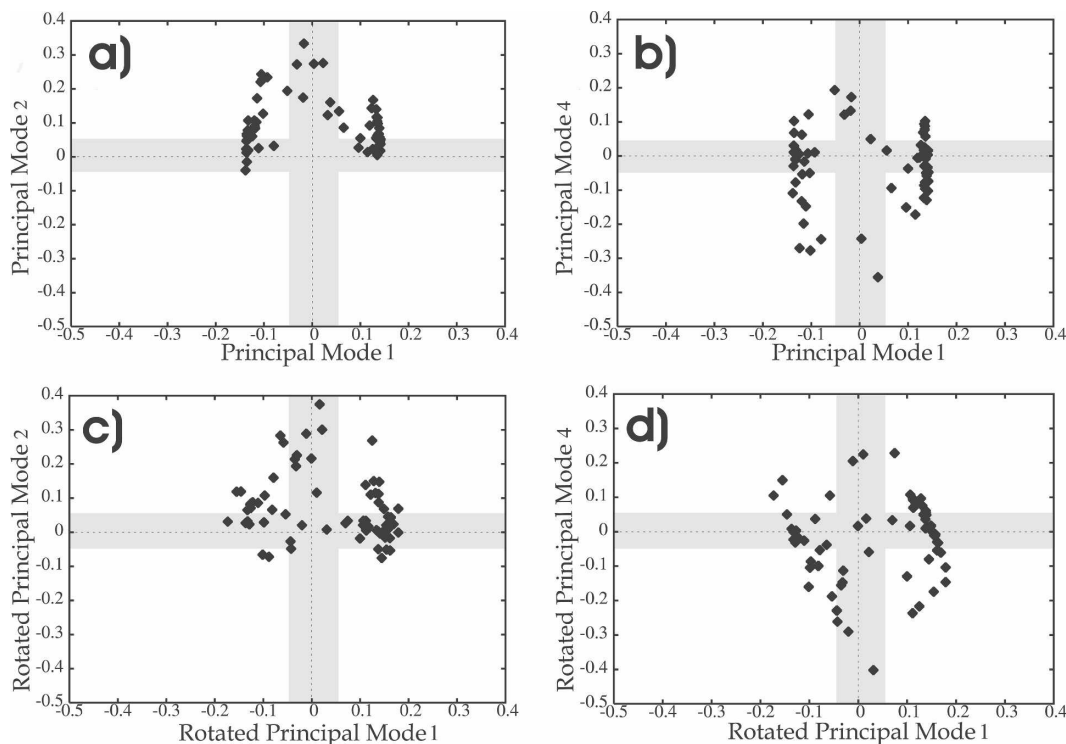


FIG. 9. Dispersion diagrams of principal mode 1 vs (a) mode 2 and (b) mode 4, and rotated mode 1 vs (c) rotated mode 2 and (d) rotated mode 4.

between the western and eastern mountains, where the rotated mode indicates perhaps some clearer tilt to the north (south). The second mode (Fig. 10b) is also similar to its unrotated analog (Fig. 6b), but with the vectors presenting a clearer alignment along the eastward (westward) direction in the mountain stations and with weaker vectors (smaller loadings) in the valleys. The third rotated principal mode (Fig. 10c) presents meridional orientation of the vectors, which show a southward (northward) sense in the eastern areas of the CFN, and zonal orientation with a westward (eastward) sense in western areas. In this case the pattern suggests a clearer behavior of the wind flow bordering the western and eastern mountain *obstacles* from some north or northeast direction. A further analysis of the synoptic conditions related to these patterns would be convenient for understanding the regional behavior as a result of the interaction of large-scale dynamics with topography. This is beyond the purposes of the regionalization at this point, but some discussion along these lines will be introduced in the last section. The fourth rotated principal mode (Fig. 10d) shows channeling flows along the northern valleys, and weak vectors in the Ebro Valley and the mountain stations. In comparison with the last unrotated vectors (Fig. 6) it distinctly highlights the channeling of wind in the northern valleys.

After a visual inspection of the loading maps (Fig. 10) a critical value for the vector modules can be defined in such a way that stations that exceed the threshold will define wind regions. As can be observed in Fig. 10, the critical value must be chosen carefully in order to avoid too much overlapping between either subregions or stations left ungrouped. Several critical values were tested, and finally the compromise was solved by adopting a critical module value of 0.175, which delimits the most consistent subregions. The regionalization obtained can be observed in Fig. 11a. The first subregion corresponds to the Ebro Valley (circles in Fig. 11a), the second is mainly defined by the mountain stations (squares in Fig. 11a), the third groups stations lined up in a north–south direction from the inner northern valleys up to the Bidasoa mountains and beyond (diamonds in Fig. 11a), and the fourth is fundamentally shaped by the northern valley sites (triangles in Fig. 11a). This wind regionalization is very similar to that obtained with the CA method (Fig. 6c). However, it clearly groups the sites labeled with diamonds (region 3) as a whole region overlapping with the northern valley sites, instead of the various small groups provided by the CA. Moreover, this regionalization creates an extended group of mountain stations instead of the very specific individual high mountain sites identified by the

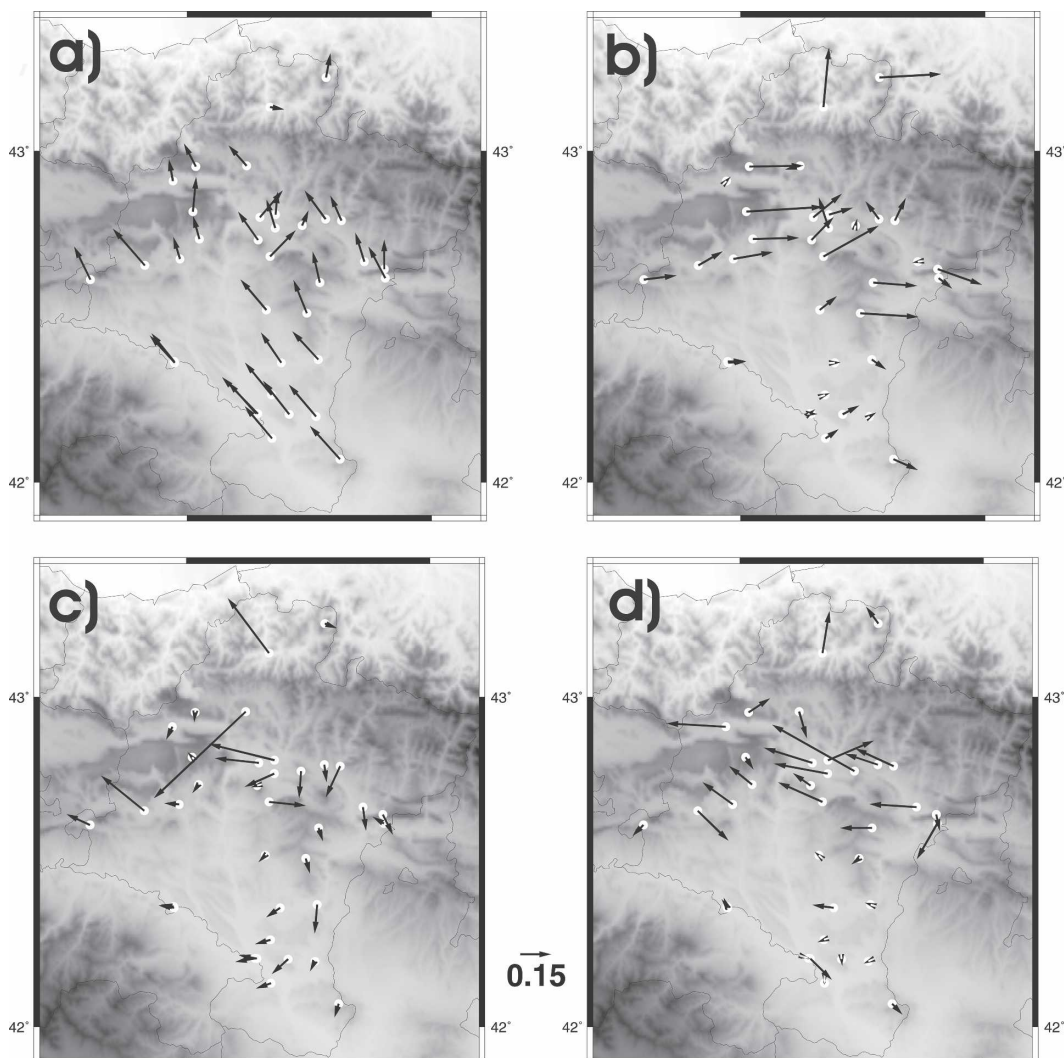


FIG. 10. Loading maps of the four rotated principal modes with the varimax technique. The tails of the vectors are placed at the observational sites. See Table 2 and Fig. 12 for complementary information.

CA regionalization method. However, this methodology has the drawback that there are two stations—10 and 16 (Fig. 1)—that were not assigned to any group. When the critical value was reduced in an attempt to include them into a group, then the subregions overlapped too much. Station 16 has the highest load module (0.168) for the Ebro Valley subregion, which agrees with the assignment resulting from the CA regionalization; however, it also has a high load for the mountain region (0.155). Station 10 has the highest load modules for the Ebro Valley subregion (0.171) and the northern valleys subregion (0.169). These types of situations could be ameliorated with further improvements in the temporal length and spatial coverage of the dataset, which would allow for capturing better signals that are faintly represented in the data.

c. Temporal variability

The second regionalization method, based on the rotation of the selected vectorial PCA modes, makes it possible to examine the temporal wind variability of each subregion by analyzing the time series of the rotated scores. For illustrative purposes, the 20-day moving-average filter outputs of the time series of scores are displayed (Fig. 12). Each score presents different behavior showing the distinct wind variability in the formed subregions. However, the rotation of the principal modes involves the loss of the independence (zero correlation) property of the scores' time series and therefore, part of the variance accounted for by one mode could also be explained by other modes (Preisendorfer 1988). The correlation among the rotated scores'

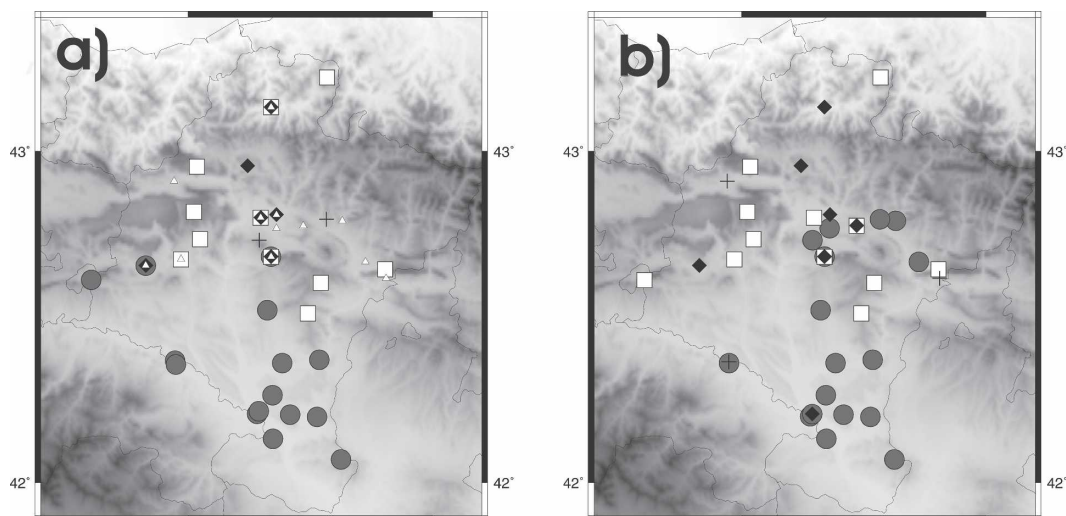


FIG. 11. Wind regionalization obtained with the rotation of the (a) four and (b) three most important principal component modes. The crosses represent the unclassified stations. The other symbols are defined in the text.

time series can be observed in Table 2. The correlation of 0.79 between score 1 (Ebro Valley) and 4 (northern valleys) implies similar wind variability in these subregions. Indeed, the regionalization reached retaining three principal modes (Fig. 11b) basically groups the stations from these two subregions. However, the fourth subregion in Fig. 11a is a well-defined cluster using the CA regionalization methodology (Fig. 8c), and is fundamentally constituted by stations that share a characteristic terrain feature, namely, location in the northern valleys. Thus, the decision was made to keep this region as a different group on the basis of the results with the CA method and the distinct topographical character.

The time series of scores of each mode are compared with the mean wind components of each wind region in Fig. 13. This allows for illustration of the meaning of the time evolution of the principal components (scores) in

relation to changes in the zonal and meridional components in each region. The changes in the scores' time series for each mode match in some regions with those of the zonal wind component, those of the meridional component, or with both. A common feature is that the four subregions present a very similar evolution of their mean meridional component, revealing that the variability of this component is uniform over all of the CFN. However, the mean zonal component shows different variability and is the actual factor that seems to play a role in distinguishing subregions. The similarity of the time series of scores with the zonal or meridional components can be understood if the orientation of vectors in the loading maps is taken into account (Fig. 10). For instance, the Ebro Valley subregion shows a similar evolution of the time series of scores and both the series of the zonal ($r = -0.93$) and meridional ($r = 0.99$) winds (Fig. 13a). This similar variability of both

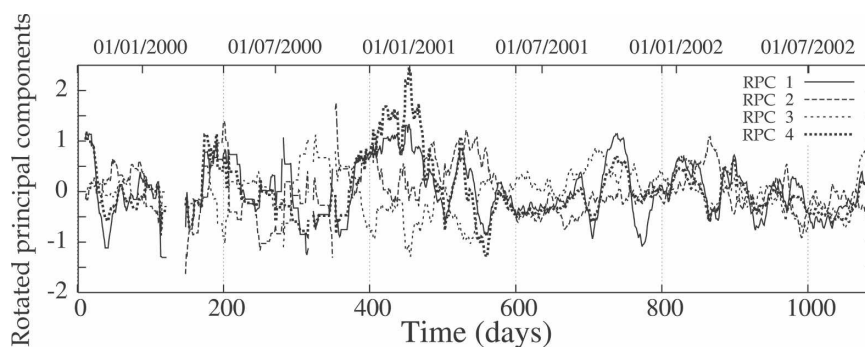


FIG. 12. The 20-day moving-average filter outputs of the time series of scores after varimax rotation are shown.

TABLE 2. Correlation of the rotated scores.

	Score 1	Score 2	Score 3	Score 4
Score 1	1.00	-0.25	0.06	0.79
Score 2	-0.25	1.00	-0.29	-0.41
Score 3	0.06	-0.29	1.00	0.09
Score 4	0.79	-0.41	0.09	1.00

wind components seems to be associated with the northwest–southeast orientation of the Ebro Valley and the channeling along it (Fig. 10a). This is also apparent in the fourth subregion (the northern valleys, see Fig. 13d), which also presents a northwest–southeast orientation of the valleys (Fig. 10d), with the meridional and zonal components similar to the score ($r = -0.84$ and $r = 0.77$, respectively). The mountain subregion evolves in agreement ($r = 0.94$) with the mean zonal component of the subregion (Fig. 13b) in concordance with the zonal orientation of the vectors in its loading map (Fig. 10b). Last, the scores in region 3 also present similar changes ($r = -0.66$) to the zonal mean

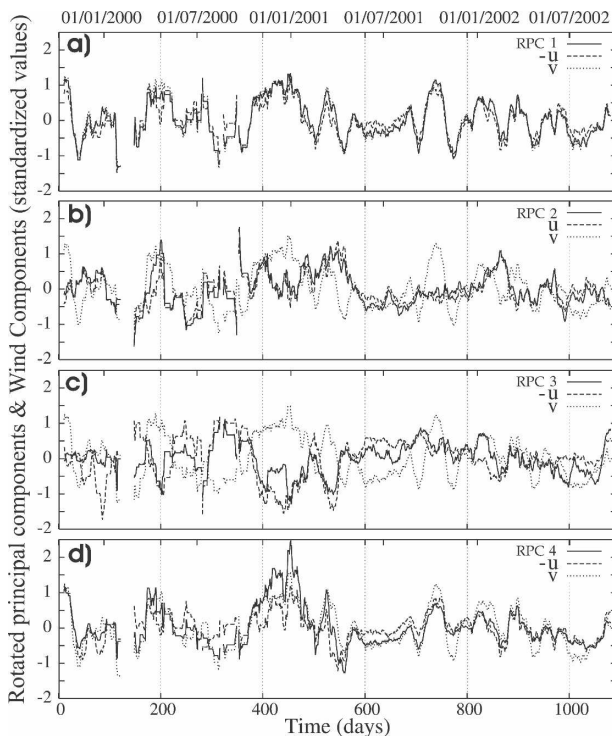


FIG. 13. The 20-day moving-average filter outputs of the time series of scores (solid lines) and the corresponding filtered u (dashed) and v (dotted) standardized mean wind components of the corresponding wind region as defined by Fig. 11a: (a) Ebro Valley (region 1), (b) mountain stations (region 2), (c) north-to-south-oriented stations (region 3), and (d) northern valleys (region 4).

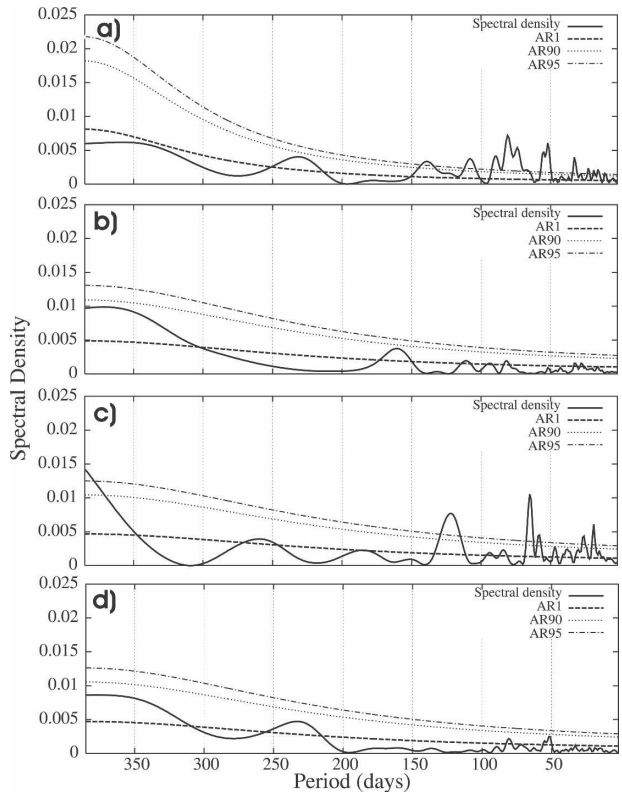


FIG. 14. Sample power spectra of the four principal component scores after varimax rotation (solid lines), their first-order autoregressive process spectra (dashed lines), and the 90% (dotted lines) and 95% (dashed-dotted lines) confidence limits. The spectra represent the wind variability in each of the following wind regions: (a) Ebro Valley (region 1), (b) mountain stations (region 2), (c) north-to-south-oriented stations (region 3), and (d) northern valleys (region 4).

wind component (Fig. 13c) resulting from the zonal orientation of the vectors with the highest loads at the stations that define the subregion in its loading map (Fig. 10c).

1) SPECTRAL ANALYSIS

Normalized spectra of the rotated scores provide a complementary understanding of wind temporal variability in each subregion (Fig. 14). The four subregions show wide spectral bands at low frequencies, although they are not significant in comparison with a red-noise autoregressive (AR1) process. It is worth noting that the time span of the series in Fig. 10 is only of about 3 yr, and thus of limited extension for significantly resolving the annual cycle. These low-frequency bands accumulate the largest portion of variance in all subregions. In addition, the various subregions accumulate variance over specific intervals centered at periods that coincide with harmonics of the annual cycle. However, few of

these are of significance in comparison with an AR1. This is the case of the Ebro Valley (region 1) and the north-to-south-oriented sites (region 3), which display significant portions of variance at frequencies ranging between 2 and 4 months. Spectra of the Ebro Valley and the northern valleys subregion (Figs. 14a,d, respectively) turn out to be very similar, as should be expected from their correlation (Table 2). However, the northern valley sites do not seem to receive AR1-significant contributions at high frequencies.

For a more complete analysis of wind variability at low frequencies, the standardized anomalies of daily wind time series from the original extended dataset can be projected onto the eigenvector of each subregion in order to reach longer time series than those in the rotated scores (769 days). Projections were performed for the days with more than 50% of the measurements available, and thus a total of 2169 days were used. This set spans over a period of more than 7 yr, from February 1995 up to September 2002. For each subregion, both the original scores and the time series obtained from projection are very similar in the overlapping parts, but the latter cover a longer time span, allowing for increased spectral resolution at lower frequencies. Spectra for the projected time series are displayed in Fig. 15. Overall, they show similar behavior as that of the score spectra (Fig. 14), with a gain in resolution and a suggestion of the presence of significant variability at yearly time scales relative to an AR1 process. As in the case of Fig. 14, only the sites in the Ebro Valley (Fig. 14a) and the north-to-south-oriented stations (Fig. 14c) show significant contributions at higher frequencies. Figures 14 and 15 suggest that the main difference between the variability of the first (Ebro Valley) and fourth (northern valleys) subregions is the apparent lack of significant contributions to high-frequency variability in the latter.

2) LARGE-SCALE INFLUENCE

Last, it can be argued that the regional variations described herein stem from the interaction of large-scale dynamics with regional topographic features and that it is of interest to explain the temporal changes in the behavior of each region on the basis of this interaction. As stated before, a thorough analysis hereon can be provided through calculation of a composite map of the standardized mean sea level pressure selecting the days on which the wind field at the surface level is characteristic of the pattern of each mode or region. These characteristic days are selected as those that present a maximum in the time series of scores (rotated principal components). The maximum ensures that the prevailing flows are those defined by the loading map

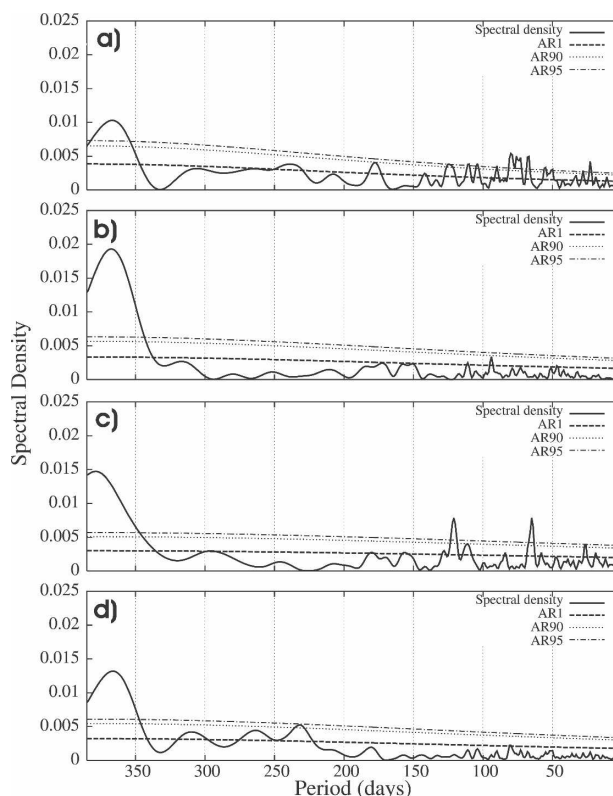


FIG. 15. As in Fig. 14, but for the standardized projections of the daily wind fields (with more than 50% of the data available) over the eigenvectors.

of the subregion (Fig. 10) resulting from the dominance of its score over the others. The composite maps of the sea level pressure for the five highest values of scores at each subregion can be observed in Fig. 16. Variations in the number of the selected days used to produce a composite do not essentially lead to different maps. The Ebro Valley subregion presents in its positive (negative) phase a negative (positive) pressure anomaly in northwestern Spain and a positive (negative) pressure anomaly in the west of Italy (Fig. 16a). This generates a northeastward (southwestward) geostrophic wind over the CFN, which turns northward (southward) because of the ageostrophic balance. This flow is channeled up (down) the Ebro Valley (Fig. 10a) and is favored by the strong pressure gradient along the valley axis that is created by the positive and negative anomaly maxima. Region 2 (mountains) shows a negative (positive) pressure anomaly in the north of France and a positive (negative) one in the west of the African coast (Fig. 16b). This develops a southwestward (northwestward) geostrophic flow, which turns eastward (westward) as a consequence of the ageostrophic balance. There is a small pressure gradient along the Ebro Valley that contributes to the weak winds on the valley (Fig. 10b). The

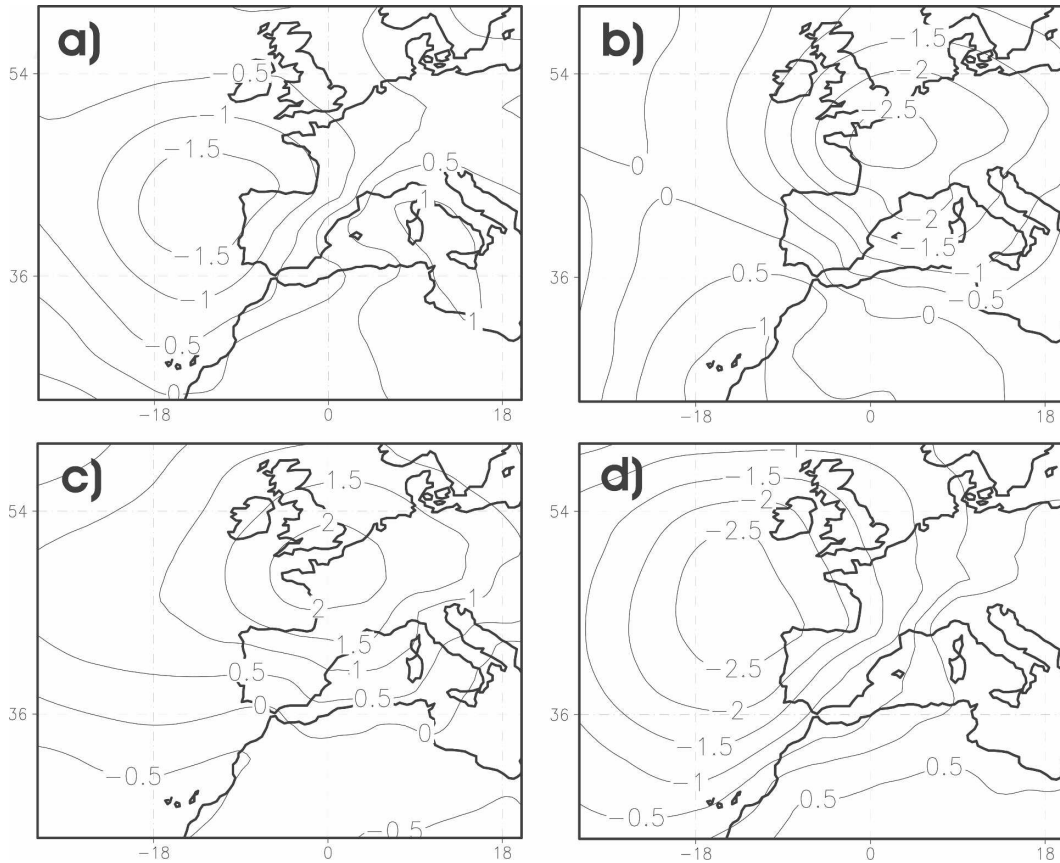


FIG. 16. Composite (averaged) maps of the standardized daily sea level pressure taken from the 40-yr European Centre for Medium-Range Weather Forecasts reanalysis data (Simmons and Gibson 2000) for the five maximum value scores of each subregion: (a) Ebro Valley (region 1), (b) mountain stations (region 2), (c) north-to-south-oriented stations (region 3), and (d) northern valleys (region 4).

north-to-south-oriented stations (region 3) present a dominant positive (negative) pressure anomaly in the north of the CFN (Fig. 16c). This anomaly generates westward or southwestward (eastward or northeastward) ageostrophic winds over the CFN (Fig. 10c). The northern valleys (region 4) present a strong negative (positive) pressure anomaly over northwestern Spain (Fig. 16d), which should generate similar winds over the CFN as the composite map of the Ebro Valley subregion (Fig. 16a). The structure of both composites is very similar, with larger negative anomalies over northwestern Spain and positive anomalies over the Mediterranean area, displaced to the south in Fig. 16d relative to Fig. 16a. The similarity of both structures is not surprising because the corresponding time series of scores were correlated ($r = 0.79$, see Table 2). The associated regional maps (Figs. 10a,d) present differences in the loadings over the Ebro Valley, which are much smaller in region 4 (Fig. 10d). These regional differences cannot be explained on the basis of the large-scale composites in Fig. 16.

5. Conclusions

Daily wind variability over a complex terrain region was analyzed. Two different methodologies based on PCA were employed to make a regionalization into subregions with similar time variability. The first approach groups together stations with similar loads by CA producing a hard regionalization where the stations belong to only one subregion (Fig. 8c). The second approach rotates the principal modes and forms subregions, allowing for some degree of overlapping (Fig. 11a). Each methodology presents its own advantages and drawbacks, and the use of both schemes allows for a good degree of robustness in conclusions; both produce an equivalent array of groups, providing wind regions in accordance with the topographic features of the terrain. The main difference between the two methodologies is that the one based on CA generates an area with a few small subregions that are joined by the rotating methodology. The CA tends to form very specific subregions, while the one based on the rotation

forms more generic groups. However, this regionalization that rotates the principal modes leaves two stations unclassified. A higher density and coverage of stations over the region would allow for a better identification of smaller subgroups, and hence it would lead to a better characterization of the wind over the CFN.

The variability of the meridional wind is very similar in all of the subgroups, and it is the zonal wind component that mostly contributes to the characterization of the wind variability in each subregion. On the basis of this statement, it can be argued that a scalar PCA-based approach applied to the zonal component would lead to similar results. While this is essentially true, such an approach would exclude the vectorial information, and thus the possibility of highlighting patterns of distinct wind direction over the CFN (e.g., channeling along the valleys). The wind spectrum of each subregion was analyzed, revealing the dominance of the annual cycle in all of them. Two subregions display significant variability at higher frequencies (2–4 months) in comparison with an AR1 process.

This wind regionalization will be used to validate mesoscale model simulations over the region. Because the simulations represent spatially filtered conditions and the observations could be affected by local effects that fall beyond model resolution, their direct comparison may be somewhat problematic. With this regionalization, wind simulations/predictions can be evaluated at each wind region, instead of at every measurement site, gaining significance in the validation procedure. This will lead to a better knowledge of model performance over the different wind regions and to potential improvement of its accuracy.

With a broader perspective, the regionalization could be useful for analyzing the dispersion of pollutants over the region, which would be different in the distinct subregions because of their different wind variability. Thus, the regionalization would be useful to evaluate impacts of existing factories or future ones, helping with the industrial development of the region. The regionalization can also prove to be useful for potentially spreading the meteorological network, helping to locate the new stations, or relocating those already existing. It would be useful also for quality-control purposes because the stations in the same subregion should present similar wind variability. Moreover, it can be employed to gain a better understanding of the relationship between synoptic-scale motions and surface flows over the region, and of the physical processes associated with wind circulations over the area.

Acknowledgments. We thank the Sección de Evaluación de Recursos Agrarios del Departamento de Ag-

ricultura, Ganadería y Alimentación of the Navarra Government for providing us with the wind dataset used in this study and the ECMWF for the free access to the ERA-40 data. We also thank Drs. M. Montoya and C. Raible for useful discussions, suggestions, and comments during this work, as well as Prof. M. Cornide for providing a first version of the code to calculate the spectral densities. The authors are indebted to the three reviewers for their comments, which helped to improve the quality of the original manuscript considerably. This work was partially funded by Project CGL2005-06966-C07/CLI. JFGR was supported by a Ramón y Cajal fellowship.

REFERENCES

- Anderberg, M. R., 1973: *Cluster Analysis for Applications*. Academic Press, 359 pp.
- Bärring, L., 1988: Regionalization of daily rainfall in Kenya by means of common factor analysis. *J. Climatol.*, **8**, 371–389.
- Belsere, E. P., 1988: Rhythms of a variable star. *Sky and Telescope*, Vol. 76, September, p. 288.
- Bloomfield, P., 1976: *Fourier Analysis of Time Series: An Introduction*. Wiley, 258 pp.
- Blumen, W., Ed., 1990: *Atmospheric Processes over Complex Terrain*. Meteor. Monogr., No. 45, Amer. Meteor. Soc., 323 pp.
- Bonell, M., and G. Sumner, 1992: Autumn and winter daily precipitation areas in Wales, 1982–1983 to 1986–1987. *Int. J. Climatol.*, **12**, 77–102.
- Cattell, R. B., 1966: The scree test for the number of factors. *Multivariate Behav. Res.*, **1**, 245–276.
- Cheng, E. D., 1998: Macroscopic extreme wind regionalization. *J. Wind Eng. Ind. Aerodyn.*, **77–78**, 13–21.
- Comrie, A. C., and E. C. Glenn, 1998: Principal components-based regionalization of precipitation regimes across the southwest United States and northern Mexico, with an application to monsoon precipitation variability. *Climate Res.*, **10**, 201–215.
- Deeming, T. J., 1975: Fourier analysis with unequally-spaced data. *Astrophys. Space Sci.*, **36**, 137–158.
- DeGaetano, A. T., 1997: A quality-control routine for hourly wind observations. *J. Atmos. Oceanic Technol.*, **14**, 308–317.
- Dyer, T. G. J., 1975: The assignment of rainfall stations into homogeneous groups: An application of principal component analysis. *Quart. J. Roy. Meteor. Soc.*, **101**, 1005–1013.
- Fovell, R. G., and M.-Y. C. Fovell, 1993: Climate zones of the conterminous United States defined using cluster analysis. *J. Climate*, **6**, 2103–2135.
- Graybeal, D. Y., 2006: Relationships among daily mean and maximum wind speeds, with application to data quality assurance. *Int. J. Climatol.*, **26**, 29–43.
- Green, M., L. O. Myrup, and R. G. Flocchini, 1992: A method for classification of wind field patterns and its application to Southern California. *Int. J. Climatol.*, **12**, 111–135.
- Gregory, S., 1975: On the delimitation of regional patterns of recent climatic fluctuations. *Weather*, **30**, 276–288.
- Hardy, D., and J. J. Walton, 1978: Principal components analysis of vector wind measurements. *J. Appl. Meteor.*, **17**, 1153–1162.

- Johnson, S. C., 1967: Hierarchical clustering schemes. *Psychometrika*, **32**, 241–254.
- Kaihatu, J. M., R. A. Handler, G. O. Marmorino, and L. K. Shay, 1998: Empirical orthogonal function analysis of ocean surface currents using complex and real-vector methods. *J. Atmos. Oceanic Technol.*, **15**, 927–941.
- Kaufmann, P., and R. O. Weber, 1996: Classification of mesoscale wind fields in the MISTRAL field experiment. *J. Appl. Meteor.*, **35**, 1963–1979.
- , and —, 1998: Directional correlation coefficient for channeled flow and application to wind data over complex terrain. *J. Atmos. Oceanic Technol.*, **15**, 89–97.
- , and C. D. Whiteman, 1999: Cluster-analysis classification of wintertime wind patterns in the Grand Canyon region. *J. Appl. Meteor.*, **38**, 1131–1147.
- Klink, K., and C. J. Willmott, 1989: Principal components of the surface wind field in the United States: A comparison of analyses based upon wind velocity, direction, and speed. *Int. J. Climatol.*, **9**, 293–308.
- Ludwig, F. L., J. Horel, and C. D. Whiteman, 2004: Using EOF analysis to identify important surface wind patterns in mountain valleys. *J. Appl. Meteor.*, **43**, 969–983.
- Mass, C. F., and Y.-H. Kuo, 1998: Regional real-time numerical weather prediction: Current status and future potential. *Bull. Amer. Meteor. Soc.*, **79**, 253–263.
- McGowan, H. A., and A. P. Sturman, 1996: Interacting multiscale wind systems within an alpine basin, Lake Tekapo, New Zealand. *Meteor. Atmos. Phys.*, **58**, 165–177.
- Meek, D. W., and J. L. Hatfield, 1994: Data quality checking for single station meteorological databases. *Agric. For. Meteorol.*, **69**, 85–109.
- Milligan, G. W., 1980: An examination of the effect of six types of error perturbation on fifteen clustering algorithms. *Psychometrika*, **45**, 325–342.
- Preisendorfer, R. W., 1988: *Principal Component Analysis in Meteorology and Oceanography*. Elsevier, 425 pp.
- Richman, M. B., 1986: Rotation of principal components. *Int. J. Climatol.*, **6**, 293–335.
- Rife, D. L., C. A. Davis, Y. Liu, and T. T. Warner, 2004: Predictability of low-level winds by mesoscale meteorological models. *Mon. Wea. Rev.*, **132**, 2553–2569.
- Romero, R., C. Ramis, J. A. Guijarro, and G. Sumner, 1999a: Daily rainfall affinity areas in Mediterranean Spain. *Int. J. Climatol.*, **19**, 557–578.
- , G. Sumner, C. Ramis, and A. Genovés, 1999b: A classification of the atmospheric circulation patterns producing significant daily rainfall in the Spanish Mediterranean area. *Int. J. Climatol.*, **19**, 765–785.
- Simmons, A. J., and J. K. Gibson, 2000: The ERA-40 Project Plan. ECMWF ERA-40 Project Rep. Series 1, Reading, United Kingdom, 63 pp.
- Sotillo, M. G., C. Ramis, R. Romero, S. Alonso, and V. Homar, 2003: Role of orography in the spatial distribution of precipitation over the Spanish Mediterranean zone. *Climate Res.*, **23**, 247–261.
- Steinacker, R., and Coauthors, 2006: A mesoscale data analysis and downscaling method over complex terrain. *Mon. Wea. Rev.*, **134**, 2758–2771.
- Stewart, J. Q., C. D. Whiteman, W. J. Steenburgh, and X. Bian, 2002: A climatological study of thermally driven wind systems of the U.S. Intermountain West. *Bull. Amer. Meteor. Soc.*, **83**, 699–708.
- Stooksbury, D. E., and P. J. Michaels, 1991: Cluster analysis of southeastern U.S. climate stations. *Theor. Appl. Climatol.*, **44**, 143–150.
- von Storch, H., 1995: Inconsistencies at the interface of climate impact studies and global climate research. *Meteor. Z.*, **4**, 72–80.
- , and F. W. Zwiers, 1999: *Statistical Analysis in Climate Research*. Cambridge University Press, 484 pp.
- Weber, R. O., and P. Kaufmann, 1995: Automated classification scheme for wind fields. *J. Appl. Meteor.*, **34**, 1133–1141.
- White, D., M. Richman, and B. Yarnal, 1991: Climate regionalization and rotation of principal components. *Int. J. Climatol.*, **11**, 1–25.
- Whiteman, C. D., 2000: *Mountain Meteorology: Fundamentals and Applications*. Oxford University Press, 355 pp.
- , and J. C. Doran, 1993: The relationship between overlying synoptic-scale flows and winds within a valley. *J. Appl. Meteor.*, **32**, 1669–1682.
- Wilks, D. S., 1995: *Statistical Methods in the Atmospheric Sciences: An Introduction*. Academic Press, 467 pp.
- Xoplaki, E., J. F. González-Rouco, J. Luterbacher, and H. Wanner, 2004: Wet season Mediterranean precipitation variability: Influence of large-scale dynamics and trends. *Climate Dyn.*, **23**, 63–78.